***Article Type:*** Systematic review

***Title:*** Concordance between core needle biopsy and surgical excision specimens for Ki-67 in breast cancer – a systematic review of the literature

***Short running title:*** Ki-67 concordance between core and surgical samples

***Authors:*** Jahnavi Kalvala[1], Ruth M Parks[1], Andrew R Green[1], Kwok-Leung Cheung[1]

***Institutions and Affiliations:***

[1]Nottingham Breast Cancer Research Centre, School of Medicine, University of Nottingham, UK

***Corresponding author:***

Professor Kwok-Leung Cheung

School of Medicine, University of Nottingham

Royal Derby Hospital Centre

Uttoxeter Road, Derby, DE22 3DT, UK

E-mail: kl.cheung@nottingham.ac.uk

Tel: +44(0)1332 724881

***Declaration of Interest***

The authors have no conflicts of interest to declare

***Word count***

3215

**Abstract**

*Aims*

The biomarkers oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor (HER)2 are routinely measured in patients with breast cancer with international consensus on how they should be interpreted. There is evidence to support use of other biomarkers to give more detailed predictive and prognostic information. Ki-67, is one example and measures the proliferative activity of cancer cells. It is important this can be done at diagnosis of breast cancer for patients who do not have initial surgical treatment (mainly older women) and those receiving neoadjuvant therapies.

*Methods*

A systematic review was performed to assess concordance of measurement of Ki-67 between core needle biopsy (CNB) samples and surgical excision (SE) samples in patients with invasive breast cancer. MEDLINE and Embase databases were searched. Studies were eligible if performed within the last 10 years; included quantitative measurement of Ki-67 in both CNB and SE sample with no prior breast cancer treatment; measured concordance between two samples; had full-text available.

*Results*

A total of 22 studies including 5982 paired CNB and SE samples on which Ki-67 was measured were appraised. Overall, there did appear to be concordance, however, reliability was unclear. Where given, the Cohen's Kappa Coefficient (k) of correlation between samples, ranged from 0.261–0.712. Concordance rate between CNB and SE where measured as a percentage, had a range from 70.3%-92.7%

*Conclusions*

Assessment of level of concordance of Ki-67 between CNB and SE samples is hampered by different methodologies. International consensus on Ki-67 measurement is urgently needed.

**Keywords*:*** core needle biopsy, surgical excision, breast cancer, Ki-67, Ki67, concordance

**Introduction**

Breast cancer is the most common female malignancy in both the developing and developed world and is the primary cause of death among women globally (1). Risk of breast cancer increases with age (2); the number of older women living with breast cancer will quadruple by 2040 (3, 4).

Core needle biopsy (CNB) is the preferred pathological method for breast cancer diagnosis compared to fine-needle aspiration cytology or surgical excision (SE) (5). The tissue removed by CNB gives information regarding tumour type, grade and expression of biomarkers such as oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor (HER)2. Measurement of these biomarkers guides therapy, as well as providing predictive and prognostic information and has become standard practice in breast cancer care.

There is growing evidence to suggest that there are other biological markers, in addition to ER, PR and HER2, which may provide more detailed predictive and prognostic information; Ki-67 is one such marker. Ki-67 is a cell-cycle specific antigen used to measure the proliferative activity of different tumour cells and is an important marker for predicting both prognosis and response to specific treatments (6-9).

Ki-67 expression in primary breast tumours has been shown to be a significant prognostic factor for overall survival (OS) (10) and is predictive of response to neoadjuvant chemotherapy (NACT) (11). Although used less commonly, there is also a role for Ki-67 monitoring in neoadjuvant endocrine therapy (NAET), especially in more recent times in the the context of the COVID pandemic, where there has been reduced surgical capacity (12). Despite this, international controversy remains regarding methodology of Ki-67 measurement, such as definition of ‘high’ versus ‘low’ expression or staining patterns (13). A

review on the subject by Urruticoechea et al (14) included 17 studies and found statistically significant association between Ki-67 and prognosis, but the cut-off for Ki-67 overexpression varied from 1% to 28.6%.

Despite these uncertainties, some associations already recommend routine measurement of Ki-67. A joint committee of the Spanish Society of Medical Oncology and the Spanish Society of Pathology (15) recommend Ki-67 to determine proliferative activity of breast cancer, however, they acknowledge that there is no absolute agreement regarding cut-off points.

A further issue is that few studies have compared the concordance rates for Ki-67 measured in CNB and SE specimens, unlike for ER, PR and HER2, for which there is a plethora of data (16-21) confirming concordance between SE and CNB measurements. By primarily measuring Ki-67 in SE specimens, the lack of data for its concordance rates could exclude a large group of women who do not undergo primary surgical treatment (mainly older women) and patients having NACT, in both clinical and reseach settings.

The aim of this paper is to systematically review the literature to determine whether there is concordance of expression of Ki-67 between CNB and SE samples, in patients receiving no prior therapy for breast cancer.

**Methods**

*Search strategy*

The OVID interface was used to search MEDLINE and Embase databases on 17th December 2020. The search was limited to those studies published within the past 10 years (1st September 2010 – 17th December 2020), those published in English language and with full text available.

Titles and abstracts were searched using the following terms: (breast cancer) AND (needle OR biopsy) AND (surgery OR surgical OR excision) AND (Ki-67 OR Ki67). Duplicate publications were excluded.

Articles were screened by two independent researchers (RP and JK) in two stages: screening of titles and abstracts, which was followed by the retrieval and screening of full-text articles. Discrepancies were resolved by discussion.

Inclusion criteria were as follows:

- Studies that included quantitative data for concordance rates for Ki-67 between SE samples and CNB samples
- Patients with invasive breast cancer, having received no prior therapy for breast cancer
- Studies with full-text available

Exclusion criteria were as follows :

- If patients received neoadjuvant treatment between CNB and SE sampling
- No quantitative measures of concordance were made
- Unable to distinguish results between in situ and invasive cases

- Failure to fulfil inclusion criteria or breast cancer not discussed
- Restricted access to study report/data/full-text

*Data extraction*

All data was extracted directly from the study text.

The following variables were extracted from the included studies when available: year of publication; country; type of study; number of paired samples; patient information including pathology of breast cancer for CNB samples and biomarker status; size of section cut from CNB and SE; antibody used to test Ki-67; definition and method of how Ki-67 was scored; Ki-67 cut-off value or range; how the scoring was performed; if the testing was performed in a hospital pathology department or a research laboratory and if it was retrospectively or prospectively performed; guidelines used for Ki-67 testing; statistical test performed to determine concordance; concordance between CNB and SE, concordance based on ER, PR and HER2 status where this information was available; interpretation of findings.

Methodological quality of full-text papers was assessed using the Reporting Recommendations for Tumour Marker Prognostic Studies (REMARK) criteria (22) and in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (23). Assessment of level of evidence, which is generally used to appraise clinical trials, was not relevant in this review, which was primarily concerned with laboratory measurement of biomarkers.

Meta-analysis was not performed due to the heterogeneity of included studies and lack of individual patient data available in each study.

## Results

*Summary*

Details of the literature search and study selection are shown in Figure 1.

A total of 22 studies met the search criteria and were included in this review. These studies included a total of 5982 paired CNB and SE breast cancer samples on which Ki-67 was measured.

Of note, all of the included studies involved female participants only.

*Reporting standard*

The results of REMARK assessment of full-text papers are given in Supplementary File 1. All papers achieved at least 65% (13 out of a maximum of 20 points) of the REMARK criteria for reporting of a study including a biomarker. The most frequently missed criterion is number 9 in relation to sample size, specified effect size, and target power.

*General characteristics of the studies*

The characteristics of the included 22 studies are presented in Table 1.

The majority of the studies were performed in Asia (11, 50%) (20, 24-33) or Europe (9, 41%) (34-42). Only one (4.5%) study was from South America (43), and one (4.5%) from North America (44). Most of the studies were retrospective in design, or used formalin-fixed paraffin embedded samples, (15, 68%), apart from two (9%), which were prospective and used fresh samples (33, 40). The design was unknown in five studies (23%) (28, 39, 41, 42, 44).

The number of paired samples included per study varied greatly, from nine (42) to 1219 (26), with an average of 271 paired samples per study. The most common histological type was

Invasive Carcinoma of No Specific Type (NST), with the majority of samples being ER/PR positive and HER2 negative.

*Methodology of the studies*

The methodology of the studies included are summarised in Table 2.

Most studies retrieved and tested samples in a hospital pathology department (13, 59%), whereas others did this in a research laboratory (4, 18%). Two studies (9%) conducted their experiments in both hospital and research laboratories (35, 40) and this information is not given in three studies (14%) (24, 28, 44).

Thirteen studies mentioned what laboratory staining instrument was used for immunohistochemistry (59%). Of these, six studies (46%) used the system by Ventana (20, 28, 30, 33, 34, 44) and four (31%) used the stainer by Dako (36, 38-40). The other manufacturers used in one study each were Leica Biosystems (27), Roche (43), and Ariol (1, 4.5%) (31).

A total of 18 studies (82%) provided data on what histological methods were used to measure Ki-67, which was immunochemistry in all cases.

Eight studies (36%) provided information on thickness of section cut, which was 3 - 5 µm in all cases. The size of SE sections was provided in only two studies (9%) (20, 28), where it was 1 cm.

A total of 18 studies (82%) described that they used the percentage of positively nuclear stained cells to count Ki-67 (20, 24, 25, 27-31, 33, 35-37, 39-44).

Ten studies (45%) described what method they used to count Ki-67 in samples. Seven of the ten studies (25, 27, 29, 30, 35, 36, 39) (70%) manually counted cells stained positive for Ki-67,

while two studies (24, 41) used an automated system (20%) and another (20) used both manual counting and an automated system (10%).

Ten studies also described how they defined scoring of cells. Seven of them counted only

Ten studies also described how they defined scoring of cells. Seven of them
counted only hotspot areas (20, 24, 27, 35-37, 39, 41, 44), one study counted hotspot and negative areas (25), and two studies counted hotspot and cold areas (36, 39).

Five studies (23%) (24, 31, 38, 39, 41) reported the incubation time of the sample with the Ki-67 antibody, which was between 20-30 minutes. Fixation time with formalin was reported in seven studies (32%) (20, 25, 27, 28, 31, 38, 44) and this varied greatly from 6-72 hours (27, 44).

A total of 14 studies (64%) used the Ki-67 clone MIB-1. Of these, 11 studies (50%) used the clone manufactured by Dako (24, 26, 29-31, 33, 36, 38-41).

The other Ki-67 clones used were 30-9 from Ventana (4, 18%) (28, 34, 35, 44) and K2 from Leica Biosystems (1, 4.5%) (27). Three studies (14%) did not provide this information (25, 32, 43).

A total of nine studies (41%) used 20% as the cut-off for 'high' expression of Ki-67 (20, 26, 30, 34-36, 38, 41, 43). Most studies (10, 45%) had more than one pathologist to score Ki-67 in the samples.

The studies also used different international guidelines including the 2011 St Gallen Consensus (45), quoted by seven studies (32%) (25, 26, 28, 32, 35, 36, 38) and the

International Ki67 in Breast Cancer Working Group (IWG) 2011 recommendations (46), quoted by two studies (9%) (20, 40).

*Rates of concordance*

The results of the studies are summarised in Table 3.

A total of nine studies (41%) used Cohen's Kappa Coefficient (k) as the statistical test to measure correlation between samples, with the value for k ranging from 0.261 – 0.712 (20, 26, 28, 30, 33, 34, 36-38). A total of ten studies (45%) measured the concordance rate between CNB and SE as a percentage, with this ranging from 70.3% - 92.7% (20, 26-28, 30, 32-34, 36, 37). Two studies (9%) (30, 40) reported median and mean Ki-67 scores in the context of ER, PR, and HER2 status, and four studies (18%) (24, 25, 29, 33) analysed whether these clinicopathological factors were associated with discordance.

**Discussion**

From the 22 studies identified in this review, it appears that there is concordance on Ki-67 measurement between CNB and SE samples, however, this cannot be quantified with the present level of data largely due to heterogeneity of the studies included.

*Reporting standard*

On average, the included studies met 16 out of a maximum of 20 points on the REMARK Criteria. The most frequently missed criterion was in relation to providing a rationale for sample size, in addition to providing the target power and effect size. In an explanation of the REMARK criteria by Altman et al (47), it is recognised that sample size receives little attention generally, due to the fact that most studies use pre-existing specimen collections or data sets, like the majority of the studies in this paper. Thus, the basis of sample size calculation is less clear and often not done.

*General characteristics of the studies*

As mentioned previously, there is discrepancy worldwide in the methodology or measurement and analysis of Ki-67. The updated recommendations from the International Ki-67 in Breast Cancer Working Group (48) have worked towards addressing some of these issues, but note that current clinical utility of Ki-67 in breast cancer care, remains limited. This present systematic review therefore is of timely interest and the first of its kind.

Most studies were retrospective in design, which is less desirable than a prospective study, which is preferable due to more accurate data collection and less variations between different centres (47). The problem retrospective studies pose is the lower quality of data as information collected retrospectively may be incomplete and may not be collected in a standardised fashion. Studies that presented no concordance data in this current review were excluded.

*Methodology of the studies*

Technical Differences

Most studies were either performed in a hospital pathology department or in a research laboratory. This information is vital as there may be a difference in testing standards, as there is no quality assurance scheme in place to measure Ki-67 in a way that is reproducible by different laboratories. (46). One potential solution to combat this issue is to have national centres that deal with Ki-67 measurements, thereby assuring that this data is standardised and comparable with other studies.

A majority of studies used the immunohistochemical staining device from Ventana, followed closely by Dako. As a potential source of inter-laboratory variation (49, 50), this is an important factor to consider. There is currently no gold-standard staining device

Formatted: Font: Not Italic, Underline

recommended (48), but quality assurance schemes should be in place to ensure the quality of staining remains constant across laboratories.

The antibody used by the majority of studies was clone MIB-1, specifically by the manufacturer Dako. Since the Ki-67 antigen is a nuclear protein and is preferentially expressed during all active phases of the cell cycle, the antigen is exclusively detected within the nucleus during interphase and mitosis. MIB-1 reacts with an epitope encoded by a repetitive element in the Ki-67 gene, and thus recognises native Ki-67 antigens and recombinant fragments from the Ki-67 molecule (51). It has been seen in studies that the MIB-1 antibody has a higher sensitivity than other antibodies, in addition to giving the best visual staining (52-55). This is corroborated by the fact that a majority of studies in this paper have used MIB-1 and thus this antibody is recommended, notably by the IWG 2020 Guidelines (48).

The fixation time for preparation of the samples, was also different across the studies and a number of the included studies mention this as a factor contributing to discordance (25-28). The IWG 2020 Guidelines (48) cite that Ki-67 immunohistochemistry seems to be more sensitive than other biomarkers to variabilities of fixation, and the index values may decrease with use of other fixatives, delays, or fixation times if not standardised (56).

Differences in Counting Methodolgy

Method of determining nuclear staining was a major difference between studies, and is likely to be a significant confounder in comparibility between studies (37). The IWG 2020 Guidelines (48) have also recommended standardised methodlogy for Ki-67 scoring. As a majority of the studies included have done, it mentions to count all positive invasive carcinoma cells within region in which all nuclei have been stained, and requires

Formatted: Underline

determination of percentage cells positive among total number of invasive cells. On the other hand, it also mentions that global scores, as opposed to hot spot methods, have higher reproducibility. A majority of studies, as mentioned above, used hot spot methods, which could be a potential source of concern when it comes to discordance. In terms of how Ki-67 was scored, a majority of studies had done this manually. The IWG 2020 guidelines also mention that there is no evidence to suggest that either automated score or visual scoring is superior to one another; this is unlikely to be significant confounder.

The IWG 2020 guidelines will hopefully go some way to standardising these factors, if adoption of the guidelines continues. They provide in detailed description of Ki-67 analysis in the following domains: preanalytical handling, standardised visual scoring system, participation in a quality control programme, potential as a prognostic marker.

The cut-offs used to define 'high' and 'low' expression of Ki-67 varied from 1% overall staining of a sample to 28.6% (14) and there is no international consensus for the cut-off (46). The St Gallen 2013 recommendations outlined that a standardised cut-off had not been established but a majority of the Panel voted that a threshold of >20% indicated 'high' Ki-67 status (57). The IWG 2020
consensus is that Ki67 5% or less, or 30% or more, can be used to estimate prognosis in certain groups, as outlined in the report (58), which is different from the St Gallen recommendations. The former also mentions that it may be useful to capture Ki-67 data "as a continuous percentage variable", which may be useful for future studies. According to the studies analysed in this present review, nine studies used 20% as the cut-off, which is in accordance with various studies (59-61). One solution to resolve the ambiguity around Ki-67 cut-offs could be to consider Ki-67 as a continuum variable, where the cut-off is dependent on the prognostic or predictive role given to Ki-67 in the specific study (62).

*Results of the studies*

Where given, the Cohen's Kappa Coefficient (k) of correlation between samples, ranged from 0.261 – 0.712 (20, 26, 28, 30, 33, 34, 36-38). This indicates a range of correlation from fair (0.21-0.40) to substantial (0.61-0.80).

Concordance rate between CNB and SE where measured as a percentage, again had a wide range from 70.3% - 92.7% (20, 26-28, 30, 32-34, 36, 37) and mirrors the findings from the correlation data; there does appear to be concordance between samples, but it has not been possible to measure the extent of this.

Given the recent recommendations by the IWG to determine Ki-67 on samples positive for ER and HER2, it is important to analyse whether ER, PR, or HER2 status can affect concordance. However, the majority of studies (15, 68%) did not consider this. In the studies that did analyse this, one study (25) reported that patients with PR-negative or HER2-positive tumours showed more Ki-67 elevation after biopsy, and another (33) reported that concordance was much higher in ER-negative tumours compared to ER-positive. On the other hand, one study (29) found no statistic difference in concordance in relation to these factors.

Overall, these percentage figures show variable concordance between CNB and SE specimens than ER, PR and HER2. Damodaran et al (16) showed that concordance between CNB and SE in 90 patient samples for ER, PR, and HER2 were 92%, 88%, and 78% respectively. Ensani et al (63) replicated these results in 100 samples, with concordance rates of 90%, 81% and 97.3% for ER, PR and HER2. The major difference between measurement of Ki-67 and these other biomarkers is that there are internationally accepted standards for measurement of ER, PR and HER2, which in turn affects confounding factors that influence Ki-67 results, such as tumour fixation and staining patterns.

There are several possible reasons why there is such a broad range of results. The study with the lowest correlation (28), k = 0.261 and 70.30%, relates this lack of correlation to differences in fixation time and tumour heterogeneity. Several studies report confounding factors that can affect concordance, such as tumor heterogeneity (20, 24, 28, 29, 33, 37) or poor standardisation of methodology such as staining patterns, surgical time interval, biopsy size and fixation time (24-27, 29, 37). While this lack of standardisation poses a problem when it comes to reproducing results, it also affects the outcome of studies as well. Since methodology is not standardised across laboratories, some studies (26, 27) cite this as a reason for their discordant results, inferring that since patients from different hospitals were referred to their centres for testing, CNB and SE specimens might have been treated differently.

The IWG 2020 Guidelines (48) reports

that Ki-67 can be a key prognostic factor if analytical validity is reached by carefully paying attention to preanalytical issues and calibrated standardised visual scoring. This further emphasises the need for an international consensus and standardised protocol for Ki-67 handelling.

*Limitations*

Studies with limited data were abundant but could not be used, such as conference abstracts or studies with no quantitative data on the concordance between CNB and SE.

The studies which have been included have used different antibodies and manufactures, cut-offs, methods to count Ki-67 and statistical methods, making it impossible to perform a meta-analysis on the data presented. The lack of standardisation among the methodologies brings

the reproducibility of these studies into question. This affects the reliability of these studies, especially if confounding factors such as if the tests were performed in a hospital pathology department or a research laboratory or if the sections of tissue samples are cut and stained the same way are not addressed. The urgency and need for a standardised international guideline for Ki-67 is evident, given the growing evidence behind its clinical and diagnostic use.

**Conclusions**

This review has identified that there are many studies set out to assess the level of concordance between Ki-67 measurement in CNB and SE samples, however, an overall opinion of this is not possible due to the lack of consensus on how Ki-67 should be measured and scored.

The majority of studies included in this present review use the Ki-67 clone MIB-1 and a cut-off of 20% (overall staining of cells) to define a sample as Ki-67 positive. The next step could be to use these criteria to measure Ki-67 on a large number of samples in a limited number of centres, to conclusively find a result.

**Declaration of Interest**

The authors have no conflicts of interest to declare

**Funding**

No funding or sponsorship was received for publication of this article. This review was conducted as part of RM Parks' PhD, supported by a Fellowship funded by Nottingham Hospitals Charity, UK and an Honorary Fellowship from the Royal College of Surgeons of England, UK.

**Author contributions**

| | |
|---|---|
| Conceived and designed the analysis | KLC, RP, JK |
| Collected the data | RP, JK |
| Contributed to analysis | RP, JK, AG |
| Prepared the manuscript | RP, JK |
| Edited the manuscript | All |

**References**


1. Benson JR, Jatoi I. The global breast cancer burden. Future Oncol. 2012;8(6):697-702.

2. Howlader N NA, Krapcho M, et al. SEER Cancer Statistics Review, 1975–2008. MD: National Cancer Institute. 2011.

3. Support MC. Breast cancer in older women. Key statistics. 2. 2012.

4. Organisation. WH. International Agency for Research on Cancer. . 2019.

5. al. DIe. Core Needle and Open Surgical Biopsy for Diagnosis of Breast Lesions: An Update to the 2009 Report [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US). 2014;Comparative Effectiveness Reviews, No. 139.

6. Oakman C, Bessi S, Zafarana E, Galardi F, Biganzoli L, Di Leo A. Recent advances in systemic therapy: new diagnostics and biological predictors of outcome in early breast cancer. Breast Cancer Res. 2009;11(2):205.

7. Yerushalmi R, Woods R, Ravdin PM, Hayes MM, Gelmon KA. Ki67 in breast cancer: prognostic and predictive potential. Lancet Oncol. 2010;11(2):174-83.

8. de Azambuja E, Cardoso F, de Castro G, Jr., Colozza M, Mano MS, Durbecq V, et al. Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients. Br J Cancer. 2007;96(10):1504-13.

9. Liu Y, Yin W, Yan T, Du Y, Shao Z, Lu J. The clinical significance of Ki-67 as a marker of prognostic value and chemosensitivity prediction in hormone-receptor-positive breast cancer: a meta-analysis of the published literature. Curr Med Res Opin. 2013;29(11):1453-61.

10. Nishimura R, Osako T, Nishiyama Y, Tashima R, Nakano M, Fujisue M, et al. Prognostic significance of Ki-67 index value at the primary breast tumor in recurrent breast cancer. Mol Clin Oncol. 2014;2(6):1062-8.

11. Kim KI, Lee KH, Kim TR, Chun YS, Lee TH, Park HK. Ki-67 as a predictor of response to neoadjuvant chemotherapy in breast cancer patients. J Breast Cancer. 2014;17(1):40-6.

12. Goncalves R, DeSchryver K, Ma C, Tao Y, Hoog J, Cheang M, et al. Development of a Ki-67-based clinical trial assay for neoadjuvant endocrine therapy response monitoring in breast cancer. Breast cancer research and treatment. 2017;165(2):355-64.

13. Petrelli F, Viale G, Cabiddu M, Barni S. Prognostic value of different cut-off levels of Ki-67 in breast cancer: a systematic review and meta-analysis of 64,196 patients. Breast Cancer Res Treat. 2015;153(3):477-91.

14. Urruticoechea A, Smith IE, Dowsett M. Proliferation marker Ki-67 in early breast cancer. J Clin Oncol. 2005;23(28):7212-20.

15. Colomer R, Aranda-López I, Albanell J, García-Caballero T, Ciruelos E, López-García M, et al. Biomarkers in breast cancer: A consensus statement by the Spanish Society of Medical Oncology and the Spanish Society of Pathology. Clin Transl Oncol. 2018;20(7):815-26.

16. Damodaran D. ea. A Prospective Study on Level of Concordance Between Core Needle Biopsy and Surgical Specimen for Assessing Oestrogen Receptor, Progesterone Receptor, and Her2/Neu Receptor Status in Carcinoma Breast and Its Implications on Treatment Decisions. Indian Journal of Surgical Oncology. 2020;11.3:446-50.

17. Espina V.A. ea. Biomarkers for predicting cancer in women with a suspicious mammogram. Annual Meeting of the American Society of Clinical Oncology, ASCO 2020. 2020;38.15.

18. Parks RMA, M.; Syed, B.M.; Green, A.R.; Ellis, I.O.; Cheung, K.-L. . Biology of Oestrogen-Receptor Positive Primary Breast Cancer in Older Women with Utilisation of Core Needle Biopsy Samples and Correlation with Clinical Outcome. Cancers. 2020;12:2067.

19. Kovacs A. ea. Comparison of immunohistochemically assessed ER, PGR, HER2, and Ki-67 breast cancer biomarker status with the automated STRAT4 RT-qPCR-based diagnostic platform-a retrospective study on 100 breast carcinomas (core biopsies & matching surgical specimens). Virchows Archiv. 2019;475(S72).

20. Jeong YS, Kang J, Lee J, Yoo TK, Kim SH, Lee A. Analysis of the molecular subtypes of preoperative core needle biopsy and surgical specimens in invasive breast cancer. J Pathol Transl Med. 2020;54(1):87-94.

21. Tamaki K, Sasano H, Ishida T, Miyashita M, Takeda M, Amari M, et al. Comparison of core needle biopsy (CNB) and surgical specimens for accurate preoperative evaluation of ER, PgR and HER2 status of breast cancer patients. Cancer Sci. 2010;101(9):2074-9.

22. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK). JNCI: Journal of the National Cancer Institute. 2005;97(16):1180-4.

23. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1-.

24. Ahn S, Lee J, Cho M-S, Park S, Sung SH. Evaluation of Ki-67 Index in Core Needle Biopsies and Matched Breast Cancer Surgical Specimens. Archives of Pathology & Laboratory Medicine. 2017;142(3):364-8.

25. Xie L, Li X, Wang Q, Zhou J, Shen J, Luo L, et al. Effects of core needle biopsy and subsequent neoadjuvant chemotherapy on molecular alterations and outcome in breast cancer. Onco Targets Ther. 2018;11:677-85.

26. You K, Park S, Ryu JM, Kim I, Lee SK, Yu J, et al. Comparison of Core Needle Biopsy and Surgical Specimens in Determining Intrinsic Biological Subtypes of Breast Cancer with Immunohistochemistry. J Breast Cancer. 2017;20(3):297-303.

27. Kombak FE, Şahin H, Mollamemişoğlu H, Önem İ, Kaya H, Buğdaycı O, et al. Concordance of immunohistochemistry between core needle biopsy and surgical resection of breast cancer. Turk J Med Sci. 2017;47(6):1791-6.

28. Chen J, Wang Z, Lv Q, Du Z, Tan Q, Zhang D, et al. Comparison of Core Needle Biopsy and Excision Specimens for the Accurate Evaluation of Breast Cancer Molecular Markers: a Report of 1003 Cases. Pathol Oncol Res. 2017;23(4):769-75.

29. Kim HS, Park S, Koo JS, Kim S, Kim JY, Nam S, et al. Risk Factors Associated with Discordant Ki-67 Levels between Preoperative Biopsy and Postoperative Surgical Specimens in Breast Cancers. PloS one. 2016;11(3):e0151054-e.

30. Chen X, Zhu S, Fei X, Garfield DH, Wu J, Huang O, et al. Surgery time interval and molecular subtype may influence Ki67 change after core needle biopsy in breast cancer patients. BMC Cancer. 2015;15(1):822.

31. Yamamoto S, Chishima T, Mastubara Y, Adachi S, Harada F, Toda Y, et al. Variability in measuring the Ki-67 labeling index in patients with breast cancer. Clin Breast Cancer. 2015;15(1):e35-9.

32. Ge W-K, Yang B, Zuo W-S, Zheng G, Dai Y-Q, Han C, et al. Evaluation of hormone receptor, human epidermal growth factor receptor-2 and Ki-67 with core needle biopsy and neoadjuvant chemotherapy effects in breast cancer patients. Thorac Cancer. 2015;6(1):64-9.

33. Chen X, Sun L, Mao Y, Zhu S, Wu J, Huang O, et al. Preoperative core needle biopsy is accurate in determining molecular subtypes in invasive breast cancer. BMC Cancer. 2013;13(1):390.

34. Robertson S, Rönnlund C, de Boniface J, Hartman J. Re-testing of predictive biomarkers on surgical breast cancer specimens is clinically relevant. Breast Cancer Res Treat. 2019;174(3):795-805.

35. Robertson S, Stålhammar G, Darai-Ramqvist E, Rantalainen M, Tobin NP, Bergh J, et al. Prognostic value of Ki67 analysed by cytology or histology in primary breast cancer. J Clin Pathol. 2018;71(9):787-94.

36. Focke C, Decker T, Diest P. Reliability of the Ki67-Labelling Index in Core Needle Biopsies of Luminal Breast Cancers is Unaffected by Biopsy Volume. Annals of surgical oncology. 2016;24.

37. Meattini I, Bicchierai G, Saieva C, De Benedetto D, Desideri I, Becherini C, et al. Impact of molecular subtypes classification concordance between preoperative core needle biopsy and surgical specimen on early breast cancer management: Single-institution experience and review of published literature. Eur J Surg Oncol. 2017;43(4):642-8.

38. Petrau C, Clatot F, Cornic M, Berghian A, Veresezan L, Callonnec F, et al. Reliability of Prognostic and Predictive Factors Evaluated by Needle Core Biopsies of Large Breast Invasive Tumors. American Journal of Clinical Pathology. 2015;144(4):555-62.

39. Knutsvik G, Stefansson IM, Aziz S, Arnes J, Eide J, Collett K, et al. Evaluation of Ki67 expression across distinct categories of breast cancer specimens: a population-based study of matched surgical specimens, core needle biopsies and tissue microarrays. PloS one. 2014;9(11):e112121-e.

40. Gandini S, Guerrieri-Gonzaga A, Pruneri G, Serrano D, Cazzaniga M, Lazzeroni M, et al. Association of molecular subtypes with Ki-67 changes in untreated breast cancer patients undergoing pre-surgical trials. Ann Oncol. 2014;25(3):618-23.

41. Romero Q, Bendahl P-O, Klintman M, Loman N, Ingvar C, Rydén L, et al. Ki67 proliferation in core biopsies versus surgical samples - a model for neo-adjuvant breast cancer studies. BMC Cancer. 2011;11(1):341.

42. Martin L-A, Davies GLS, Weigel MT, Betambeau N, Hills MJ, Salter J, et al. Pre-surgical study of the biological effects of the selective cyclo-oxygenase-2 inhibitor celecoxib in patients with primary breast cancer. Breast Cancer Research and Treatment. 2010;123(3):829-36.

43. Rey-Vargas L, Mejía-Henao JC, Sanabria-Salas MC, Serrano-Gomez SJ. Effect of neoadjuvant therapy on breast cancer biomarker profile. BMC Cancer. 2020;20(1):675.

44. Clark BZ, Onisko A, Assylbekova B, Li X, Bhargava R, Dabbs DJ. Breast cancer global tumor biomarkers: a quality assurance study of intratumoral heterogeneity. Modern Pathology. 2019;32(3):354-66.

45. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ. Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. Ann Oncol. 2011;22(8):1736-47.

46. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. J Natl Cancer Inst. 2011;103(22):1656-64.

47. Altman DG, McShane LM, Sauerbrei W, Taube SE. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK): explanation and elaboration. PLoS medicine. 2012;9(5):e1001216-e.

48. Nielsen TO, Leung SCY, Rimm DL, Dodson A, Acs B, Badve S, et al. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. JNCI: Journal of the National Cancer Institute. 2020.

49. Raap M, Ließem S, Rüschoff J, Fisseler-Eckhoff A, Reiner A, Dirnhofer S, et al. Quality assurance trials for Ki67 assessment in pathology. Virchows Archiv. 2017;471(4):501-8.

50. Røge R, Nielsen S, Riber-Hansen R, Vyberg M. Impact of Primary Antibody Clone, Format, and Stainer Platform on Ki67 Proliferation Indices in Breast Carcinomas. Appl Immunohistochem Mol Morphol. 2019;27(10):732-9.

51. Key G, Becker MH, Baron B, Duchrow M, Schlüter C, Flad HD, et al. New Ki-67-equivalent murine monoclonal antibodies (MIB 1-3) generated against bacterially expressed parts of the Ki-67 cDNA containing three 62 base pair repetitive elements encoding for the Ki-67 epitope. Lab Invest. 1993;68(6):629-36.

52. Boccon-Gibod LA. Pathological evaluation of renal tumors in children: International Society of Pediatric Oncology approach. Pediatric and developmental pathology. 1998;1(3):243-8.

53. Youssef EM, Matsuda T, Takada N, Osugi H, Higashino M, Kinoshita H, et al. Prognostic significance of the MIB-1 proliferation index for patients with squamous cell carcinoma of the esophagus. Cancer. 1995;76(3):358-66.

54. Lindboe CF, Torp SH. Comparison of Ki-67 equivalent antibodies. J Clin Pathol. 2002;55(6):467-71.

55. Barbareschi M, Girlando S, Mauri FM, Forti S, Eccher C, Mauri FA, et al. Quantitative growth fraction evaluation with MIB1 and Ki67 antibodies in breast carcinomas. Am J Clin Pathol. 1994;102(2):171-5.

56. Arima N, Nishimura R, Osako T, Nishiyama Y, Fujisue M, Okumura Y, et al. The importance of tissue handling of surgically removed breast cancer for an accurate assessment of the Ki-67 index. J Clin Pathol. 2016;69(3):255-9.

57. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Annals of oncology : official journal of the European Society for Medical Oncology. 2013;24(9):2206-23.

58. Nielsen TO, Leung SCY, Rimm DL, Dodson A, Acs B, Badve S, et al. Assessment of Ki67 in Breast Cancer: Updated Recommendations From the International Ki67 in Breast Cancer Working Group. JNCI: Journal of the National Cancer Institute. 2021;113(7):808-19.

59. Bustreo S, Osella-Abate S, Cassoni P, Donadio M, Airoldi M, Pedani F, et al. Optimal Ki67 cut-off for luminal breast cancer prognostic evaluation: a large case series study with a long-term follow-up. Breast Cancer Res Treat. 2016;157(2):363-71.

60. Tashima R, Nishimura R, Osako T, Nishiyama Y, Okumura Y, Nakano M, et al. Evaluation of an Optimal Cut-Off Point for the Ki-67 Index as a Prognostic Factor in Primary Breast Cancer: A Retrospective Study. PLoS One. 2015;10(7):e0119565.

61. Goldhirsch A, Winer EP, Coates AS, Gelber RD, Piccart-Gebhart M, Thürlimann B, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. Ann Oncol. 2013;24(9):2206-23.

62. Denkert C, von Minckwitz G. Reply to Ki67 in breast cancer: a useful prognostic marker! Ann Oncol. 2014;25(2):542-3.

63. Ensani F, Omranipour R, Jahanzad I, Jafari A, Nafarzadeh S, Aminishakib P. The Core Needle and Surgical Biopsy Concordance to Detect Estrogen, Progesterone, and Her-2 Receptors in Breast Cancer: A Comparative Study. Iran J Pathol. 2017;12(3):202-8.

64. Ekholm M, Grabau D, Bendahl PO, Bergh J, Elmberger G, Olsson H, et al. Highly reproducible results of breast cancer biomarkers when analysed in accordance with national guidelines - a Swedish survey with central re-assessment. Acta Oncol. 2015;54(7):1040-8.

65. Weidner N, Moore DH, 2nd, Vartanian R. Correlation of Ki-67 antigen expression with mitotic figure index and tumor grade in breast carcinomas using the novel "paraffin"-reactive MIB1 antibody. Hum Pathol. 1994;25(4):337-42.

**Table 1:** Characteristics of the 22 studies included in the systematic review

| # | Author and date | Country | Number of paired samples | Mean age (range) in years | Histological type on CNB | ER Status | PR Status | HER2 Status | Tumor grade on CNB | Type of Study |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | L Rey-Vargas et al (43), July 2020 | Colombia | 61 | - | - | +ve 46 (75.4%) | +ve 46 (75.40%) | +ve 14 (23%) | - | Retrospective |
| 2 | Jeong YS et al (20), Nov 2019 | South Korea | 629 | Median 53 (23- 89) | IDC NST 504 (80%), ILC 45 (7.2%) Mucinous 24 (3.8%), Medullary 13 (2.1%), Metaplastic 8 (1.3%), Minor type 35 (5.6%) | - | - | - | I 143 (22.7%), II 286 (45.5%), III 200 (31.8%) | Retrospective |
| 3 | S Robertson et al (34), Jan 2019 | Sweden | 526 | Median 65 (26-97) | IDC NST 316(60.1%), ILC 99(18.8%), Mucinous 15 (2.9%), Tubular 10(1.9%), Papillary 4(0.8%), Mixed 5(1%), U 77(14.6%) | - | - | - | I 65 (12.4%), II 269 (51.1%), III 189 (35.9%), U 3 (0.6%) | Retrospective |
| 4 | BZ Clark et al (44), Oct 2018 | USA | 82 | 60 (33 - 92) | IDC 89 (89%), ILC 6 (6%), Mucinous 4 (4%), Metaplastic 1 (1%) | - | - | - | I 13 (13%), II 54 (54%), III 33 (33%) | - |
| 5 | S Robertson et al (35), | Sweden | 369 | 62 | IDC NST 324 (87.8%), ILC 33 (8.9%), Mucinous 11 (2.9%), Tubular 2 (0.54%), Papillar 2 | - | - | - | I 52 (14.1%), II 177 (48.0%), III 154 (41.7%) | Retrospective |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Sept 2018 | | | | (0.54%), Medullary 2 (0.54%), Mixed 5 (1.36%), Other subtype 6 (1.6%), U 1 (0.27%) | | | | | |
| 6 | S Ahn et al (24), ~~March~~ Mar 2018 | South Korea | 89 | - | IDC 86 (96.6%), Metaplastic 1 (1.12%), Papillary 1 (1.12%), Mucinous 1 (1.12%) | +ve 60 (67.4%) | +ve 55 (61.8%) | +-ve 14~~75~~ (16~~84~~%) | - | Prospective |
| 7 | L Xie et al (25), ~~February~~ Feb 2018 | China | 123 | - | - | 0 33 (26.8%) I 9 (7.3%), II 23(18.7%), III 48 (39%) | 0 50 (40.7%), I 21 (17.1%), II 22(17.9%, III 30 (24.4%), U 30(24.4%) | +-ve 25~~98~~ (20~~79.7~~%) | - | Retrospective |
| 8 | K You et al (26), ~~September~~ Sept 2017 | South Korea | 1219 | Median 49.5 (24-86) | - | - | - | - | I 317 (24.1%), II 590 (44.9%), III 407 (31.0%) | Retrospective |
| 9 | FE Kombak et al (27), ~~August~~ Aug 2017 | Turkey | 236 | 52.3 (22-84) | IDC 245 (86.5%), ILC 8 (2.8%), Mixed and other 31 (12.5%) | +ve 248 (87.3%) | +ve 229 (80.6%) | 0 159 (65.4%), I 16 (6.6%), II 22 (9%), | - | Retrospective |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | III 46 (18.9%) | | |
| 10 | J Chen et al (28), ~~January~~ Jan 2017 | China | 696 | Median 47 (16-89) | IDC 876 (87.3%), ILC 61 (6.1%), Mucinous 14 (1.4%), Medullary 13 (1.3%), Apocrine 4 (0.4%), Papillary microcarcinoma 16 (1.6%), Tubular 11 (1.1%), Adenoid cystic 2 (0.2%), Scirrbous 1 (0.1%) | - | - | - | I 37 (3.7%), II 346 (34.5%), III 620 (61.8%) | Retrospective |
| 11 | CM Focke et al (36), ~~December~~ Dec 2016 | Germany | 170 | - | - | - | - | - | - | - |
| 12 | I Meattini et al (37), ~~October~~ Oct 2016 | Italy | 101 | Median57.5 (29-86) | IDC 64(63.4%), ILC 13(12.8%), IDC + ILC 14(13.8%), Metaplastic 3 (3%), Colloid 2(2%), Apocrine 2(2%), Other 3(3%) | +ve 80 (3=79.2%) | +ve 86 (85.1%) | - | - | Retrospective |
| 13 | HS Kim et al (29), ~~March~~ Mar 2016 | South Korea | 310 | -- | - | - | - | - | - | Retrospective |
| 14 | X Chen et al (30), | China | 276 | 56.6 (24-91) | IDC 246 (89.1%), ILC 12 (4.3%), Others 18 (6.5%) | +ve 214 (77.5%) | +ve 163 (59.1%) | +-ve 59~~217~~ | I 6 (2.2%), II 134 (48.6%), | Retrospective |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Oct 2015 | | | | | | | (21.3 %) | III 90 (32.6%), U 46 (16.7%) | |
| 15 | S Yamamoto et al (31), ~~February~~ Feb 2015 | Japan | 87 | Median 57 (21-87) | IDC 80 (91.9%), ILC 2 (2.3%), Other 5 (5.7%) | - | - | - | I 34 (39.1%), II 41 (47.1%), III 12 (13.8%) | Retrospective |
| 16 | WK Ge et al (32), ~~January~~ Jan 2015 | China | 82 | Median 49 (27-70) | - | - | - | - | - | Retrospective |
| 17 | C Petrau et al (38), ~~January~~ Jan 2015 | France | 163 | 63 (30-93) | IDC 133 (82%), ILC 22(13%), Mucinous 3(2%), Mixed 2(1%), Other type 2(1%), U 1(1%) | +ve 135 (83%) | +ve 94 (58%) | 0 or I+ 131 (80%), II+ 14 (9%), III + 18 (11%) | I 26 (16%), II 98(60%), III 38(23%), U 1(1%) | Retrospective |
| 18 | G Knutsvik et al (39), ~~November~~ Nov 2014 | Norway | 137 | - | IDC 447(83.7%), ILC 55(10.3%), Tubular 8(1.5%), Mucinous 16(3.0%), Medullary 4(0.7%), U 4(0.7%) | +ve 451 (84.5%) | +ve 337 (70.6%) | +-ve 7146~~3~~ (13.3~~86.7~~ %) | I 218 (40.8%), II 226 (42.3%), III 90 (16.9%) | Retrospective |
| 19 | S Gandini et al (40), ~~March~~ Mar 2014 | Italy | 269 | Median 51 | - | - | - | - | - | - |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 20 | X Chen et al (33), ~~August~~ Aug 2013 | China | 298 | Median 54 (21-91) | IDC 260 (87.2%), ILC 12(4%),<br>Mixed 11(3.7%), Others 15(5%) | +ve<br>231 (77.5%) | +ve<br>178 (59.7%) | +-ve<br>65~~233~~ (21.8~~78.2~~ %) | I 6 (2%),<br>II 138 (46.3%),<br>III 97 (32.6%),<br>U 57 (19.1%) | Prospective |
| 21 | Q Romero et al (41), ~~August~~ Aug 2011 | Sweden | 50 | - | IDC 44(88%),<br>ILC 5(10%), Mucinous 1(2%) | +ve<br>45 (90%) | +ve<br>37 (74%) | - | I 16(32%),<br>II 19(38%),<br>III 15(30%) | Retrospective |
| 22 | LA Martin et al (42), ~~August~~ Aug 2010 | UK | 9 | 71 (59-89) | - | +ve<br>7 (77%) | +ve<br>5 (56%) | +ve<br>1 (11%) | II 3 (33%),<br>III 6 (67%) | - |

**Abbreviations:** Ductal Carcinoma in Situ (DCIS); Invasive Ductal Carcinoma (IDC); Invasive Lobular Carcinoma (ILC); Invasive Ductal Carcinoma of Non-Specific Type (IDC NST); Estrogen Receptor (ER); Progesterone Receptor (PR); Human Epidermal Growth Receptor 2 (HER2); Positive (+ve); Negative (-ve); Unknown (U); Core Needle Biopsy (CNB); Study Number (#)

**Table 2:** Methodology of the 22 studies included in the systematic review

| # | Laboratory instrument used for IHC | Thickness of section | Antibody used to test Ki-67 (clone, manufacturer, dilution) | Counting Method ~~dology~~ manual or automated | ~~Definition of Scoring~~Counting method hotspot or global | Incubation Time with Antibody | Fixation time | Ki-67 Cut-off for determining 'high' expression | Who did the test/how was it done? - Scoring | Where was measurement of Ki-67 performed? | Guidelines used for Ki-67 Testing |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Roche Benchmark XT | CNB - 3 µm | - | - | - | - | - | 20% | Single pathologist | ~~Research Laboratory~~ RL | - |
| ~~2~~ | ~~Ventana BenchMark Ultra~~ | ~~—~~ | ~~MIB-1, Dako, 1:100~~ | | | | | ~~20%~~ | ~~—~~ | ~~Research Laboratory~~ | ~~According to manufacturer's recommendations (Lab Vision Corporation)~~ |

Formatted Table

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | Ventana Benchmark XT ~~slide stainer~~ | CNB - 4 µm; SE - 0.5 - 1 cm | MIB-1, Ventana, prediluted | Manual (eyeball) and automated(Ventana) | ~~Manual~~ Hotspots | - | 8 - 24 hours | 20% | 4 pathologists | ~~Hospital Pathology Department~~ HPD | ~~International Ki67 in Breast Cancer Working Group~~IWG 2011 (46) |
| 3~~4~~ | Ventana BenchMark ULTRA ~~Staining Module, Ventana Medical Systems~~ | - | 30-9, Ventana | - | - | - | - | 20% | - | ~~Research Laboratory~~ RL | ~~St. Gallen consensus meeting in 2013~~SG 2013 (57)~~(63)~~ and current Swedish guidelines (64) |
| 4 | Ventana | - | 30-9, Ventana | - | Hotspots ~~(10% higher than average over whole slide)~~ | - | 6 - 72 hours | 25% | - | - | - |
| 5~~6~~ | - | - | 30-9, Ventana | Manual | Hotspot~~s areas~~ | - | - | 20% | 2 pathologists | ~~Research Laboratory and Hospital Pathology~~ | - |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | RL RL and HPD | |
| 6 | - | CNB - 5 µm | MIB-1, Dako, 1:500 | Digital image analysis software (Definiens) | Hotspots ~~area at x200 magnification~~ | 25 mins | - | - | Digital Image Analysis; 1 pathologist | - | - |
| 7 | ~~Performed by the Department of Pathology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine.~~ | - | - | Manual | Hotspot and Negative areas~~Manual~~ | - | 24 hours | 14% | 2 pathologists | ~~Hospital Pathology Department~~ HPD | - |
| 8 | - | - | MIB-1, Dako | - | - | - | - | 20% | - | ~~Hospital Pathology Department~~ HPD | ~~St. Gallen's~~SG 2011 (45) |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | Bond-Max autoimmunostainer, (Leica Biosystems | CNB - 4 µm | K2, Leica Biosystems, 1:100 | Manual | Maual 0 Hotspots) | - | 6 - 72 hours | 14% | 2 pathologists | Hospital Pathology Department HPD | St. Gallen'sSG 2011 (45) |
| 10 | Ventana Benchmark System | CNB - 4 µm, SE - 1 cm | 30-9, Ventana | - | - | - | 8 - 12 hours | 14% | 2 pathologists | - | - |
| 11 | Autostainer (LINK48 Autostainer, Dako) | CNB - 4 µm | MIB-1, Dako, 1:200 | Manual | Maual - Hotspot, coldspot and an intermediate area | - | - | 20% | - | Hospital Pathology Department HPD | St. Gallen'sSG 2011 (45) |
| 12 | - | - | MIB-1, Immunotech | - | Hotspots | - | - | 15% | - | Hospital Pathology Department HPD | St. Gallen'sSG 2011 (45) |
| 13 | - | - | MIB-1, Dako | Manual | - | - | - | 14% | single pathologist | Hospital Pathology Department HPD | |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 14 | BenchMark XT, Ventana ~~Autostain System~~ | - | MIB-1, Dako | Manual | - | - | - | 20% | 2 pathologists | ~~Hospital Pathology Department~~ HPD | Department of Pathology, Severance Hospital of Yonsei University College of Medicine |
| 15<br>~~6~~ | Ariol SL-50 | - | MIB-1, Dako | - | - | 20 mins | 48 hours | 14% | - | HPD | - |
| 16 | - | - | - | - | - | - | - | 14% | 2 pathologists | ~~Hospital Pathology Department~~ HPD | - |
| 17 | Autostainer Link 48, Dako | - | MIB-1, Dako, 1:250 | - | - | 20 mins | 48 hours | 20% | 2 pathologists | ~~Hospital Pathology Department~~ HPD | ~~St. Gallen's~~SG 2011 (45) |
| 18 | DAKO autostainer | CNB - 5 μm | MIB1, Dako, 1:100 | Manual | ~~Manual~~ Hotspot and coldspot areas | 30 mins | - | - | single pathologist | ~~Research Laboratory~~ RL | ~~St. Gallen's S~~ 2011 (45) |

Formatted: Normal, No bullets or numbering

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 19 | Automated immunostainer, (Dako) | - | MIB-1, Dako, 1:50 | - | - | - | - | - | - | RL and HPD Research Laboratory and Hospital Pathology Department | Methodology used in Weidner et al (65) |
| 20 | Ventana autostain system, BenchMark XT, Ventana | - | MIB-1, Dako | - | - | - | - | 14% | 2 pathologists | Research Laboratory RL | International Ki67 in Breast Cancer Working GroupIWG 2011 (46) |
| 21 | - | CNB - 4 µm | MIB-1, Dako, 1:500 | Automated System (DAKO) | Hotspots - | 25 mins | - | 20% | 2 observers | Hospital Pathology Department HPD | - |
| 22 | - | | MIB-1 | - | - | - | - | - | 2 observers | Hospital Pathology Depar | - |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | HPD | |

**Abbreviations** Core Needle Biopsy (CNB); Surgical Excision (SE); Micrometer (μm); Centimeter (cm); Study Number (#); Research Laboratory (RL); Hospital Pathology Department (HPD); International Ki67 in Breast Cancer Working Group (IWG); St. Gallen (SG)

**Table 3:** Findings of the 22 studies included in the systematic review

| # | Statistical Test | Concordance - k value | Concordance - % | Concordance - p | Concordance - Other | Concordance based on ER/PR/HER2 Status | Interpretation of findings |
|---|---|---|---|---|---|---|---|
| 1 | | - | - | p = 0.5796 | Median values show no difference | - | No significant variations found in biomarker status or expression values. |
| 2 | Cohen's Kappa Coefficient | 0.577 | 78.70% | p < 0.001 | - | - | CNB was reasonably accurate for determining Ki-67 and molecular subtypes. |
| 3 | Cohen's Kappa Coefficient | 0.529 | 78.80% | p<0.001 | Median % in CNB = 24%, Median % in SE = 28%, | - | The agreement of CNB and paired surgical specimen in primary breast cancer is insufficient. |
| 4 | STATISTICA, p-values using Wilcoxon-matched pairs test. Coefficient of variation and | - | - | p = 0.010931 | Mean for CNB = 31%, Mean for SE = 36%. | - | Biomarker expression on either sample is representative of breast carcinoma in most cases and is appropriate for clinical decision-making. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | quartile coefficient of dispersion. | | | | | | |
| 5 | Mann-Whitney U tests, Pearson's Correlation, Spearman's rank, Cohen's Kappa | - | - | - | r = 0.639 | - | Differences in the prognostic value of Ki-67 in breast cancer depending on the diagnostic method were seen. |
| 6 | Contingency tables and y2 tests, t test | - | - | - | Median absolute difference = 3.5% (P<0.01) | In the 13 cases that showed discordance, 12 were ER+ and 3 were HER2+. | Difference in Ki-67 index between core biopsy and surgical specimens was observed. Advised to perform Ki-67 assay on both core needle biopsy and the surgical specimen. |
| 7 | Paired t-test, Wilcoxon signed-rank test | - | - | - | Mean - CNB 19.1% Mean - SE 24%, P<0.01 | Patients with PR- or HER2+ showed more obvious Ki-67 elevation | CNB samples could provide more reliable information on determination of molecular subtype than SE. |
| 8 | Cohen’s Kappa Coefficient | 0.712 | 87% | p <0.001 | - | - | CNB showed high diagnostic accuracy compared with surgical specimens, and good agreement. |
| 9 | - | - | 80.90% | - | P < 0.001; Median Ki- 67 positivity of 15% for CNBs and 20% for RSs | - | Higher number of discordant cases for Ki-67 as compared to other biomarkers. |
| 10 | Cohen’s Kappa Coefficient | 0.261 | 70.30% | - | P < 0.001 | - | Results obtained using CNB should be used cautiously when determining treatment. |
| 11 | Cohen’s Kappa Coefficient | 0.44 | 75% | - | - | - | Reliability of Ki-67 levels in CNB of luminal breast cancers is unaffected by CNB volume. |
| 12 | Cohen’s Kappa Coefficient | 0.68 | 88.10% | p = 0.0001 | - | - | Should be detected both on CNB and SE samples, especially in |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | | hormonal positive HER2 negative tumours |
| 13 | T test, Fisher's test | - | - | - | Median levels = 10%, P<0.001, Spearman's rho = 0.676 | There was no statistical difference in clinicopathological characteristics except histologic grade between cases with Ki-67 | Showed a substantial concordance. Extremely discordant Ki-67 levels may be associated with aggressive tumor biology. |
| 14 | Cohen's Kappa Coefficient, two paired samples t test. | 0.6 | 80.40% | - | Mean SE = 29.1%, Mean CNB = 26.2% P < 0.001 | In HER2+ samples – Ki-67 Median change of 5 | Ki-67 value significantly increased after CNB, which was associated with molecular subtype. |
| 15 | Pearson correlation coefficient | - | - | - | CNB - 56.3%, SE = 55.2% | - | There was a significant difference in concordance rates when using different Ki-67 cut-offs. |
| 16 | $X^2$ test, Spearman correlation coefficient | - | 92.70% | - | R = 0.842, P<0.01 | - | Neither CNB nor surgical excision samples gave highly consistent results in Ki-67 status. |
| 17 | Cohen's Kappa Coefficient | 0.55 | - | - | - | - | The correlation was weak for Ki-67. |
| 18 | Spearman rank coefficient, Bland and Altman analysis, Wilcoxon signed rank tests, Mann-Whitney U or Kruskal-Wallis H-test | - | - | p =0.56 | Median %, SE = 17%, CNB = 13%, P<0.001, P=0.001 | Ki-67 in SE (N=534) - ER+ = 16.6% ER- = 42.8% PR+ = 16.8% PR- = 26.2 HER2+ = 32.4%; HER2- = 16.8%. | There is a significant difference in tumor cell proliferation by Ki-67 across different sample categories. Our findings indicate that specimen specific cut-off values should be applied for practical use. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | Ki-67 in CNB (N=154)<br>ER+ = 11%<br>ER-=40%<br>PR+ = 11.1%<br>PR-19.3%<br>HER2+ = 18.4%;<br>HER2- = 11.7% | |
| 19 | Random Effects Model | - | - | - | Median Change was 0 or 0%, Mean Change was 2.2 +/- 9.2% and 13.58% | Median Ki-67 on HER2+ samples – CNB = 30, SE = 35.5.<br>Mean Ki-67 on HER2+ samples - CNB = 33.3, SE= 38.4 | A significant increase in Ki-67 was ascertained in HER2-positive and triple-negative tumours. |
| 20 | - | 0.545 | 79.50% | - | P < 0.001 | HER2 status had no impact on K-i67. However, Ki-67 concordance rate was higher in ER- tumors compared with ER positive tumors (92.5% vs. 76.2%, *P* = 0.003). Moreover, patients with PR -, or grade 3 tumors had a better agreement using CNB to detect Ki-67 status than those with PR + or grade | Ki-67 should be retested on SE samples in HR+/HER2- patients to accurately distinguish Luminal A from B tumours. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | 1–2 diseases, with *P* value 0.012 and 0.006, respectively. | |
| 21 | Paired t test, bland-altman analysis | - | - | p = 0.19 | mean proliferation 2.2% higher than SE | - | Importantly, the difference in concordance represents an average difference in proliferation with the core biopsies demonstrating a higher proliferation index compared to the surgical samples. |
| 22 | - | - | - | - | Mean Change = -8.1%, P = 0.24 | - | High Ki-67 levels were associated with higher tumour grade. None of the statistical methods employed show a significant difference between core biopsy and corresponding surgical sample proliferation values. |

**Abbreviations** Cohen's Kappa Coefficient (k); Spearman Correlation Coefficient (r); Percentage (%); Core Needle Biopsy (CNB); Surgical Excision (SE); Estrogen Receptor (ER); Progesterone Receptor (PR); Human Epidermal Growth Factor Receptor 2 (HER2); Positive (+ve); Negative (-ve); Hormone Receptor (HR)